

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-339322

(43)Date of publication of application : 08.12.2000

(51)Int.Cl.

G06F 17/30 G06F 13/00
G06F 17/00

(21)Application number : 11-145675

(71)Applicant : NIPPON TELEGR & TELEPH CORP
<NTT>

(22)Date of filing : 25.05.1999

(72)Inventor : INAGAKI HIROTO
TANAKA KAZUO

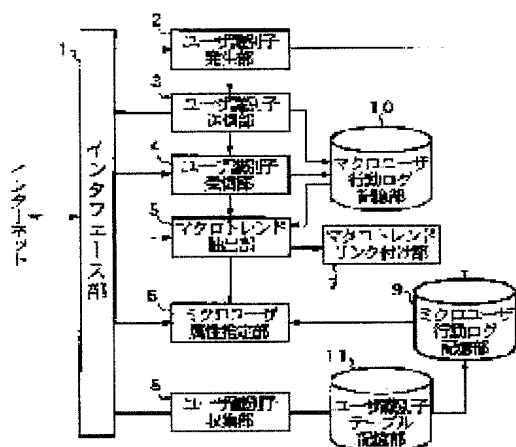
(54) METHOD AND DEVICE FOR SPECIFYING USER ATTRIBUTE AND RECORDING MEDIUM WITH ITS METHOD

(57)Abstract:

PROBLEM TO BE SOLVED: To estimate a user's attribute based on the user's behavior having a purpose clearly without the direct input of the user's attribute from a keyboard, etc., by each user.

SOLUTION: This device has a user identifier generating part 2 which generates a user identifier for identifying an accessing user, a user identifier transmitting part 3 which transmits the generated user identifier to the accessing user, a user identifier receiving part 4 that receives a user identifier transmitted from the terminal of the user when the user starts a behavior, a macro trend extracting part 5 which extracts a macro trend being the trend of a macroscopic user behavior based on the received user identifier, and a micro user attribute estimating part 6 which estimates the region, racial characteristics, sex,

interest and income to which the user having at least a user identifier belongs based on the user identifier, the macro trend, the regional characteristics and temporal characteristics of the macro trend.



対応なし、英抄

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2000-339322

(P2000-339322A)

(43) 公開日 平成12年12月8日 (2000.12.8)

(51) Int.Cl. ⁷	識別記号	F I	デマンド* (参考)	
G 0 6 F 17/30		C 0 6 F 15/403	3 4 0 A	5 B 0 4 9
13/00	3 5 4	13/00	3 5 4 D	5 B 0 7 5
17/00		15/20	N	5 B 0 8 9

審査請求 未請求 請求項の数 9 O L (全 12 頁)

(21) 出願番号 特願平11-145675

(22) 出願日 平成11年5月25日 (1999.5.25)

(71) 出願人 000004226

日本電信電話株式会社

東京都千代田区大手町二丁目3番1号

(72) 発明者 稲垣 博人

東京都新宿区西新宿三丁目19番2号 日本
電信電話株式会社内

(72) 発明者 田中 一男

東京都新宿区西新宿三丁目19番2号 日本
電信電話株式会社内

(74) 代理人 100064908

弁理士 志賀 正武

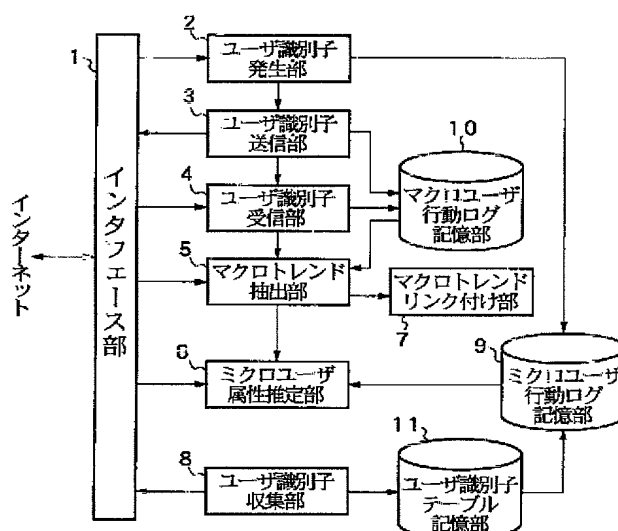
最終頁に続く

(54) 【発明の名称】 ユーザ属性特定方法、ユーザ属性特定装置、および、その方法を記録した記録媒体

(57) 【要約】

【課題】 ユーザの属性を、各ユーザが直接キーボード等から入力することなく、ユーザの明確に目的づけられた行動に基づいてユーザの属性を推定すること。

【解決手段】 アクセスしてきたユーザを識別するためのユーザ識別子を生成するユーザ識別子発生部2と、前記生成したユーザ識別子を前記アクセスしてきたユーザに送信するユーザ識別子送信部3と、前記ユーザが行動を起こした際に、該ユーザの端末から送信されてくるユーザ識別子を受信するユーザ識別子受信部4と、前記受信したユーザ識別子に基づき巨視的なユーザの行動の傾向であるマクロトレンドを抽出するマクロトレンド抽出部5と、前記ユーザ識別子と、前記マクロトレンドと、該マクロトレンドの地域的特性および時間的特性に基づき、少なくとも該ユーザ識別子を持つユーザの属する地域、民族性、性別、趣味、所得を推定するマイクロユーザ属性推定部6とを持つ。



【特許請求の範囲】

【請求項1】 ネットワークを介してアクセスしてきたユーザの属性を特定するユーザ属性特定方法において、前記アクセスしてきたユーザを識別するためのユーザ識別子を生成する第1のステップと、前記生成したユーザ識別子を前記アクセスしてきたユーザに送信する第2のステップと、前記ユーザが行動を起こした際に、該ユーザの端末から送信されてくるユーザ識別子を受信する第3のステップと、前記受信したユーザ識別子に基づき巨視的なユーザの行動の傾向であるマクロトレンドを抽出する第4のステップと、前記ユーザ識別子と、前記マクロトレンドと、該マクロトレンドの地域的特性および時間的特性に基づき、少なくとも該ユーザ識別子を持つユーザの属する地域、民族性、性別、趣味、所得を推定する第5のステップとを有することを特徴とするユーザ属性特定方法。

【請求項2】 前記第4のステップで抽出したマクロトレンドのうち、同一の嗜好をもつマクロトレンドをリンク付けする第6のステップを有することを特徴とする請求項1に記載のユーザ属性特定方法。

【請求項3】 前記ネットワークに接続されたサーバから、該サーバが前記ユーザに対して提供した識別子を収集し、前記第1のステップで生成したユーザ識別子と共に、ユーザ毎にテーブル化する第7のステップを有すること特徴とする請求項1または2に記載のユーザ属性特定方法。

【請求項4】 ネットワークを介してアクセスしてきたユーザの属性を特定するユーザ属性特定装置において、前記アクセスしてきたユーザを識別するためのユーザ識別子を生成するユーザ識別子発生部と、前記生成したユーザ識別子を前記アクセスしてきたユーザに送信するユーザ識別子送信部と、前記ユーザが行動を起こした際に、該ユーザの端末から送信されてくるユーザ識別子を受信するユーザ識別子受信部と、前記受信したユーザ識別子に基づき巨視的なユーザの行動の傾向であるマクロトレンドを抽出するマクロトレンド抽出部と、前記ユーザ識別子と、前記マクロトレンドと、該マクロトレンドの地域的特性および時間的特性に基づき、少なくとも該ユーザ識別子を持つユーザの属する地域、民族性、性別、趣味、所得を推定するユーザ属性推定部とを持つことを特徴とするユーザ属性特定装置。

【請求項5】 前記マクロトレンド抽出部により抽出されたマクロトレンドのうち、同一の嗜好をもつマクロトレンドをリンク付けするマクロトレンドリンク付け部を有することを特徴とする請求項4に記載のユーザ属性特定装置。

【請求項6】 前記ネットワークに接続されたサーバから、該サーバが前記ユーザに対して提供した識別子を収集し、前記ユーザ識別子発生部により生成されたユーザ識別子と共に、ユーザ毎にテーブル化するユーザ識別子収集部を有することを特徴とする請求項4または5に記載のユーザ属性特定装置。

【請求項7】 ネットワークを介してアクセスしてきたユーザの属性を特定するユーザ属性特定プログラムを記録したコンピュータ読み取り可能な記録媒体であって、前記アクセスしてきたユーザを識別するためのユーザ識別子を生成する第1のステップと、前記生成したユーザ識別子を前記アクセスしてきたユーザに送信する第2のステップと、前記ユーザが行動を起こした際に、該ユーザの端末から送信されてくるユーザ識別子を受信する第3のステップと、前記受信したユーザ識別子に基づき巨視的なユーザの行動の傾向であるマクロトレンドを抽出する第4のステップと、前記ユーザ識別子と、前記マクロトレンドと、該マクロトレンドの地域的特性および時間的特性に基づき、少なくとも該ユーザ識別子を持つユーザの属する地域、民族性、性別、趣味、所得を推定する第5のステップとを有するユーザ属性特定プログラムを記録した記録媒体。

【請求項8】 前記第4のステップで抽出したマクロトレンドのうち、同一の嗜好をもつマクロトレンドをリンク付けする第6のステップを有することを特徴とする請求項7に記載のユーザ属性特定プログラムを記録したコンピュータ読み取り可能な記録媒体。

【請求項9】 前記ネットワークに接続されたサーバから、該サーバが前記ユーザに対して提供した識別子を収集し、前記第1のステップで生成したユーザ識別子と共に、ユーザ毎にテーブル化する第7のステップを有することを特徴とする請求項7または8に記載のユーザ属性特定プログラムを記録したコンピュータ読み取り可能な記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、ユーザの行動からそのユーザの属性を自動的に特定するユーザ属性特定方法およびユーザ属性特定装置ならびにその記録媒体に関する。

【0002】

【従来の技術および発明が解決しようとする課題】従来、ユーザの属性は、各ユーザが直接キーボード等から入力した情報、もしくは、ユーザの視線や動作を検知してそれら視線や動作に基づいて推定することにより特定していた。しかし、ユーザの視線や行動からユーザの属性を自動的に推定する方法では、ユーザの視線や行動が、曖昧であったり明確に目的づけられたものでない場

合が多く、的確なユーザ属性の推定が難しかった。

【0003】また、インターネットでは、ブラウザ利用者の属性情報を収集する場合に、ユーザを識別するユーザ識別子として、cookieというものが使われている。これは、インターネットの端末に対して、ある特定のURLにアクセスしようとした際に、そのURLのHTTPサーバは、cookieと呼ばれるユーザ識別子を、アクセスしてきた端末に対して送信している。このcookieは、正確には、各ユーザの識別子というわけではなく、あるユーザがある端末を使用した時に付与されるユーザ識別子であり、例えば、家と仕事場とで異なる端末を使っている場合は、必ずしも同じユーザ識別子とはならない。

【0004】さらにcookieは、そのcookieが蓄積されている端末を操作した人を特定するためのものであるが、1台の端末が複数の人で利用される場合には、的確なマイクロユーザログの収集が難しい。また、cookieだけでは、違うHTTPサーバに接続すると違うサーバのcookieを取得することになってしまうこともあり、本当のユーザの属性を取得することができない等の欠点がある。よって、ホームページを開設している側では、実際にブラウジングした利用者の属性を把握することができなかった。

【0005】本発明の目的は、ユーザの属性を、各ユーザが直接キーボード等を利用して入力することなく、ユーザの明確に目的づけられた行動だけを選び出し、その明確に目的づけられた行動に基づいてユーザの属性を推定するユーザ属性特定方法およびユーザ属性特定装置ならびにその記録媒体を提供することを目的とする。

【0006】

【課題を解決するための手段】請求項1に記載の発明は、ネットワークを介してアクセスしてきたユーザの属性を特定するユーザ属性特定方法において、前記アクセスしてきたユーザを識別するためのユーザ識別子を生成する第1のステップと、前記生成したユーザ識別子を前記アクセスしてきたユーザに送信する第2のステップと、前記ユーザが行動を起こした際に、該ユーザの端末から送信されてくるユーザ識別子を受信する第3のステップと、前記受信したユーザ識別子に基づき巨視的なユーザの行動の傾向であるマクロトレンドを抽出する第4のステップと、前記ユーザ識別子と、前記マクロトレンドと、該マクロトレンドの地域的特性および時間的特性に基づき、少なくとも該ユーザ識別子を持つユーザの属する地域、民族性、性別、趣味、所得を推定する第5のステップとを有すること特徴としている。

【0007】請求項2に記載の発明は、請求項1に記載のユーザ属性特定方法において、前記第4のステップで抽出したマクロトレンドのうち、同一の嗜好をもつマクロトレンドをリンク付けする第6のステップを有すること特徴としている。

【0008】請求項3に記載の発明は、請求項1または

2に記載のユーザ属性特定方法において、前記ネットワークに接続されたサーバから、該サーバが前記ユーザに対して提供した識別子を収集し、前記第1のステップで生成したユーザ識別子と共に、ユーザ毎にテーブル化する第7のステップを有すること特徴としている。

【0009】請求項4に記載の発明は、ネットワークを介してアクセスしてきたユーザの属性を特定するユーザ属性特定装置において、前記アクセスしてきたユーザを識別するためのユーザ識別子を生成するユーザ識別子発生部と、前記生成したユーザ識別子を前記アクセスしてきたユーザに送信するユーザ識別子送信部と、前記ユーザが行動を起こした際に、該ユーザの端末から送信されてくるユーザ識別子を受信するユーザ識別子受信部と、前記受信したユーザ識別子に基づき巨視的なユーザの行動の傾向であるマクロトレンドを抽出するマクロトレンド抽出部と、前記ユーザ識別子と、前記マクロトレンドと、該マクロトレンドの地域的特性および時間的特性に基づき、少なくとも該ユーザ識別子を持つユーザの属する地域、民族性、性別、趣味、所得を推定するユーザ属性推定部とを持つことを特徴としている。

【0010】請求項5に記載の発明は、請求項4に記載のユーザ属性特定装置において、前記マクロトレンド抽出部により抽出されたマクロトレンドのうち、同一の嗜好をもつマクロトレンドをリンク付けするマクロトレンドリンク付け部を有すること特徴としている。

【0011】請求項6に記載の発明は、請求項4または5に記載のユーザ属性特定装置において、前記ネットワークに接続されたサーバから、該サーバが前記ユーザに対して提供した識別子を収集し、前記ユーザ識別子発生部により生成されたユーザ識別子と共に、ユーザ毎にテーブル化するユーザ識別子収集部を有すること特徴としている。

【0012】請求項7に記載の発明は、ネットワークを介してアクセスしてきたユーザの属性を特定するユーザ属性特定プログラムを記録したコンピュータ読み取り可能な記録媒体であって、前記アクセスしてきたユーザを識別するためのユーザ識別子を生成する第1のステップと、前記生成したユーザ識別子を前記アクセスしてきたユーザに送信する第2のステップと、前記ユーザが行動を起こした際に、該ユーザの端末から送信されてくるユーザ識別子を受信する第3のステップと、前記受信したユーザ識別子に基づき巨視的なユーザの行動の傾向であるマクロトレンドを抽出する第4のステップと、前記ユーザ識別子と、前記マクロトレンドと、該マクロトレンドの地域的特性および時間的特性に基づき、少なくとも該ユーザ識別子を持つユーザの属する地域、民族性、性別、趣味、所得を推定する第5のステップとを有するユーザ属性特定プログラムを記録したことを特徴としている。

【0013】請求項8に記載の発明は、請求項7に記載

のユーザ属性特定プログラムを記録したコンピュータ読み取り可能な記録媒体において、前記ユーザ属性特定プログラムは、前記第4のステップで抽出したマクロトレンドのうち、同一の嗜好をもつマクロトレンドをリンク付けする第6のステップを有することを特徴としている。

【0014】請求項9に記載の発明は、請求項7または8に記載のユーザ属性特定プログラムを記録したコンピュータ読み取り可能な記録媒体において、前記ユーザ属性特定プログラムは、前記ネットワークに接続されたサーバから、該サーバが前記ユーザに対して提供した識別子を収集し、前記第1のステップで生成したユーザ識別子と共に、ユーザ毎にテーブル化する第7のステップを有することを特徴としている。

【0015】本発明では、ホームページのアクセスログを分析することでユーザのアクセス行動からユーザがどのような情報に関心があり、どのような情報を求めているかを推測し、この推測からユーザ属性（ユーザのプロファイル）を推定する。

【0016】具体的には、インターネット上に存在する検索エンジンのホームページ（検索サイトともいう）において、利用者が投入した検索キーワード、および、複数表示された検索結果候補のうち、さらに詳細な内容を表示させるためにクリックすることにより選択したホームページの中で、検索結果候補の表示画面に再び戻ってくるまでの時間が長かったホームページほどその利用者の関心度が高いと推定する。

【0017】

【発明の実施の形態】次に、本発明に係るユーザ属性特定装置の一実施形態について、図面を参照して説明する。図1は本実施形態のユーザ属性特定装置の構成を示すブロック図である。なお、本実施形態のユーザ属性特定装置は、インターネット上でホームページ（特に検索エンジン等を利用しているホームページ）を開設しており、このホームページを利用するユーザの属性を特定するものである。

【0018】図1において、1はインターネットに接続されたインターフェイス部であり、インターネットを介して本実施形態のユーザ属性特定装置にアクセスしてきたユーザと、各種情報の授受を行う。2はユーザ識別子発生部であり、本実施形態のユーザ属性特定装置にアクセスしてきたユーザに対応した、個々のユーザを識別するための識別子（以下、ユーザ識別子という）を生成する。3はユーザ識別子送信部であり、ユーザ識別子発生部1が発生したユーザ識別子を、ユーザインターフェイス部1を介してそのユーザ識別子に対応するユーザへ送信する。

【0019】4はユーザ識別子受信部であり、各ユーザが行動を起こした際に当該各ユーザの識別子を、ユーザインターフェイス部1を介して受信する。5はマクロ

トレンド抽出部であり、ユーザ識別子受信部4が受信したユーザ識別子に基づいてユーザの行動の傾向を巨視的に推定する。ここで、ユーザの行動とは、例えば、ユーザがあるURLの表示をブラウザに指示したり、ホームページ検索を指示したり、あるコンテンツを見ることをサーバに要求したりすることを意味する。

【0020】6はマイクロユーザ属性推定部であり、ユーザ識別子受信部4が受信したユーザ識別子と、マクロトレンド抽出部5が推定した巨視的なユーザ行動の傾向（大規模なユーザ集合の代表的な行動。以下、マクロトレンドという）と、マクロトレンドの地域的特性や時間的特性とに基づいて、当該ユーザ識別子を持つユーザが属する地域、民族性、性別、趣味、所得等を推定する。7はマクロトレンドリンク付け部であり、同一の嗜好をもつマクロトレンドをリンクづけする。8はユーザ識別子収集部であり、各サイトからユーザ識別子のキーを取り出し、そのキーをもとに各ユーザの識別子を収集する。

【0021】9はマイクロユーザ行動ログ記憶部であり、個人個人の、または、ある特定のユーザ識別子をもつグループの、微視的なユーザの行動ログを記憶している。10はマクロユーザ行動ログ記憶部であり、個人個人ではなく、逆に、巨視的に大規模なユーザ集合の代表的な行動ログを記憶している。すなわち、例えば、あるユーザがあるURLを辿って別のページに行った場合、ユーザのアクション毎にユーザが辿ったURLがマクロユーザ行動ログとしてマクロユーザ行動ログ記憶部10に記憶される。また、上記ユーザが辿ったURLが、上記ユーザのユーザ識別子に対応づけられてマイクロユーザ行動ログとしてマイクロユーザ行動ログ記憶部9に記憶される。

【0022】これにより、マクロユーザ行動ログ記憶部10には、各ユーザが、どのURLから他のサイトに来訪し、そして別のサイトにジャンプしていったのか、また、同一サイト内の別のページにジャンプしたのかが記録される。また、マイクロユーザ行動ログ記憶部9には、各ユーザのユーザ識別子毎に、各ユーザが、どのURLから他のサイトに来訪し、そして別のサイトにジャンプしていったのか、また、同一サイト内の別のページにジャンプしたのかが記録される。11はユーザ識別子テーブル記憶部であり、各ユーザが持つすべてのユーザ識別子を収集したテーブルを記憶している。

【0023】次に上述した構成により、ユーザ属性を特定する際の動作について、図2に示すフローチャートを参照して詳細に説明する。ユーザが目的とするホームページを探索するために上記ユーザ属性特定装置にアクセスすると、まず、ステップS1において、ユーザ識別子発生部2がインターネットのcookieと同様なメカニズムで、上記アクセスしてきた端末（以下、単に端末という）にユニークなユーザ識別子を発生する。例えば、ユ

ユーザ識別子を示す変数名を Keyとした場合、ユーザ識別子は、Key=123456789となる。

【0024】ここで、変数名Key（以下、単にKeyという）は、そのユーザ識別子を端末から取り出す際のリレーショナルデータベース（以下、RDBという）のキーと同様な働きをする。また、ユニークなユーザ識別子を発生するために、RDB等の機能を利用してユニークなユーザ識別子を発生する。また、ユーザ識別子発生部2は、発生したユーザ識別子をマイクロユーザ行動ログ記憶部9へ記憶させる。

【0025】次にステップS2へ進み、ユーザ識別子送信部3は、まず、Keyを利用して、Keyで登録されているユーザ識別子があるかどうかを上記端末に問い合わせる。そして、もし、Keyで登録されているユーザ識別子がないとの応答があった場合は、ステップS3へ進み、ユーザ識別子発生部1で発生したユーザ識別子をユーザ側に送信し、上記端末の記憶装置に保存させる。また、Keyで登録されているユーザ識別子が既に存在していた場合は、ステップS4へ進み、上記端末に対しそのユーザ識別子を送信させる。

【0026】ここで、上述した動作では、まずステップS1で、上記端末に対するユーザ識別子を発生した後、ステップS2で、上記端末に対してユーザ識別子があるかどうかの問い合わせを行っているが、最初に端末に対してユーザ識別子があるかどうかの問い合わせを行い、ユーザ識別子がないとの応答があった場合にユーザ識別子を発生し、それを端末の記憶装置に保存させるようにしてもよい。

【0027】ステップS3またはS4の処理が完了すると、次にステップS5へ進み、ユーザ識別子受信部4は、ユーザが行動を起こした時に、そのユーザが使用している端末から送信されてくる当該ユーザのユーザ識別子を受信する。ここで、ユーザが起こす行動としては、ホームページ探索のための検索キーの入力行為や、ホームページ検索の結果、ユーザの端末画面に表示されるホームページのタイトルまたはURL（Uniform Resource Locator）のうち、ユーザが所望するページへジャンプするために、そのページのタイトルまたはURLを指定する行為等がある。

【0028】次に、ステップS6へ進み、マクロトレンド抽出部5およびマイクロユーザ属性推定部6により、ユーザ識別子に基づくマクロトレンドを抽出およびする。まず、マクロトレンド抽出部5によるマクロトレンドの抽出について説明する。マクロトレンド抽出部5は、マクロユーザ行動ログを蓄積、解析することによりマクロトレンドを抽出する。これは、例えば、大久保らによる、「情報関連づけ装置およびその方法」（特開平10-320419号）を用いることにより、精度良くマクロトレンドを抽出することができる。

【0029】ここで、図3を参照してマクロトレンド抽

出部5により実行されるマクロトレンドの抽出手順について説明する。まず、マクロトレンド抽出部5は、ユーザ識別子受信部4を介してユーザが目的とするホームページを検索するための検索キーが入力されたか否かを判断する（ステップSa1）。もし、検索キーが入力されなかった場合は、判断結果がNOとなって、マクロトレンド抽出部5は処理を終了する。

【0030】また、もし、検索キーが入力された場合は、判断結果がYESとなって入力された検索キーに基づいてホームページを検索し、これにより検索されたホームページのタイトルまたはURLとその内容（すなわち検索結果）をユーザの端末に一覧表示させる（ステップSa2）。そして、それと同時に計時を開始し、所定時間（例えば1～2分間）以内にユーザが検索結果として一覧表示されたページのタイトルまたはURLのうち、いずれかをクリックしたか否かを判断する（ステップSa3）。

【0031】ここで、所定時間以内に一覧表示されたページのタイトルまたはURLが何等クリックされなかった場合は、判断結果がNOとなってマクロトレンド抽出部5は処理を終了する。また、所定時間以内にいずれかのタイトルまたはURLがクリックされた場合は、判断結果がYESとなり、そのクリックされたタイトルまたはURL（検索結果）が、過去において同じユーザ識別子を持つユーザによりクリックされていたか否かを判断する（ステップSa4）。

【0032】そして、過去において同じユーザ識別子を持つユーザにより、ステップSa3でクリックされた同じタイトルまたはURLと同じものがクリックされていた場合は、判断結果がYESとなって、マクロトレンド抽出部5は処理を終了する。一方、上記ユーザが初めてクリックしたページのタイトルまたはURLであった場合は、判断結果がNOとなって、上記ユーザがステップSa1で入力された検索キーと、上記ユーザがクリックしたページのタイトルまたはURLと、当該クリックされたページが、今までにクリックされた回数の累計（すなわち、他のユーザによりクリックされた回数も含んでいる）とを、マクロユーザ行動ログとしてマクロユーザ行動ログ記憶部10に記憶する。ここで、上記累計は傾向の大きさ（トレンド強度）を示していると見なせる。

【0033】次に、マクロトレンド抽出部5は、マクロユーザ行動ログ記憶部10に記憶されているマクロユーザ行動ログを参照して、 Δ 累計/ Δ 時間の値が一定値以上であるか否かを判断する（ステップSa6）。ここで、 Δ 累計/ Δ 時間の値は、時間の経過に伴い変化する累計の度合い（トレンドの傾き）を示している。そして、 Δ 累計/ Δ 時間の値が一定値未満である場合は、判断結果がNOとなり、マクロトレンド抽出部5は処理を終了する。また、 Δ 累計/ Δ 時間の値が一定値以上である場合は、判断結果がYESとなり、所定時間が経過し

てから、今度は Δ 累計/ Δ 時間の値が一定値以下であるか否かを判断する(ステップSa7)。

【0034】そして、 Δ 累計/ Δ 時間の値が一定値を超えていれば、累計が未だピークに達していないと見なし、判断結果がNOとなってマクロトレンド抽出部5は処理を終了する。一方、所定時間経過後の Δ 累計/ Δ 時間の値が一定値以下であれば、累計がピークに達したと見なし、判断結果がYESとなって、次にステップSa3でクリックされたページの、クリック回数の累計が一定値以上であるか否かを判断する(ステップSa8)。ここで、累計が一定値未満であった場合は、判断結果がNOとなり、検索キーの内容や検索結果はマクロトレンドではないとしてマクロトレンド抽出部5は処理を終了する。

【0035】一方、累計が一定値以上であった場合は、判断結果がYESとなり、ステップSa1で入力された検索キーと、ステップSa3でクリックされた検索結果(ページのタイトルまたはURL)と、そのページのクリック累計がピークに達した時刻とを、マクロトレンドとして記憶し(ステップSa9)、再度検索キーの入力待機状態となる(ステップSa1)。

【0036】以上のような処理を行うことにより、例えば、桜の開花の時期には、“桜”に関する検索や“桜”に関係するサイトへのジャンプが増加し、図4に示すような、桜の開花の時期である3月末にピークが出現するマクロトレンド(一般的な傾向)を抽出することができる。

【0037】次に、マイクロユーザ属性推定部6によるユーザ属性の推定について説明する。マイクロユーザ属性推定部6は、ユーザ識別子受信部4が受信したユーザ識別子と、マクロトレンド抽出部5が抽出したマクロトレンドと、このマクロトレンドの地域的特性や時間的特性に基づいて、そのユーザ識別子を持つユーザの、属する地域、民族性、性別、趣味、所得等を推定する。

【0038】すなわち、マイクロユーザ属性推定部6は、マクロトレンド抽出部5により抽出されたマクロトレンドに対し、さらに地域的特性や、時間的特性に基づく“傾向”を抽出する。例えば、地域的な特性がある“トレンド”の場合は、地域的軸を拡散させ、マクロトレンドを計算する。また、時間的特性に基づく傾向の場合は、時間軸を拡散させ、マクロトレンドを再計算する。

【0039】これにより、マクロトレンド抽出部5により抽出されたマクロトレンドが、地域的トレンドや、時間的トレンドに変化する。そして、その地域的トレンドや時間的トレンドをダイナミックに解析することにより、ユーザの特性を抽出することができる。例えば、図4に示したマクロトレンドを、地域毎に分けて抽出すると、図5に示すようにいくつかのピークが出現する。なお、図5における点線は図4のマクロトレンドを示している。

【0040】これは、“桜”の開花が、時間的にも地域的にも軸があるトレンドであることを示しているのである。つまり、図5において、最初のピークは、例えば九州地方のマクロトレンドであり、次のピークは東京または本州のマクロトレンドであり、最後のピークは東北地方または北海道地方のマクロトレンドであることがわかる。

【0041】そして、このマクロトレンドを発生しているユーザ識別子を、マイクロユーザ行動ログ記憶部9から抽出することにより、あるユーザ識別子を持つユーザが、東京地方に住んでいる人か、九州地方に住んでいる人か、東北地方に住んでいる人かというユーザの地域属性を推定することができる。さらにユーザの地域属性をローカルに絞る場合は、上記のような地域的なマクロトレンドを、さらに複合的にかけて、絞り込むことにより、たとえば、あるユーザ識別子をもつユーザが、東京の杉並区にすんでおり、一人住まいで、年収がどのくらいであるかのユーザ属性を特定することが可能となる。

【0042】すなわち、複数の地域的なマクロトレンドを掛け合わせて複合した場合、それぞれのマクロトレンドが集合であると考え、その集合の和を求めることとなるので、地域をさらに的確に絞ることができる。例えば、あるマクロトレンドが関東地方のユーザを識別し、ある別のマクロトレンドが東京以北のユーザを識別するとすれば、この2種類のマクロトレンドにより、関東地方でかつ東京以北のユーザを識別することができる。この例では、東京、埼玉、茨城のユーザであることが推定される。

【0043】以下、図6に示すフローチャートを参照してマイクロユーザ属性推定部6におけるユーザ属性の推定手順について説明する。まず、マイクロユーザ属性推定部6は、マクロトレンド抽出部5と同様、ユーザ識別子受信部4を介してユーザが目的とするホームページを検索するための検索キーが入力されたか否かを判断する(ステップSb1)。もし、検索キーが入力されなかった場合は、判断結果がNOとなって、マイクロユーザ属性推定部6は処理を終了する。

【0044】一方、もし検索キーが入力された場合は、前述したように、マクロトレンド抽出部5によって、ユーザが入力した検索キーに基づいてホームページが検索され、検索結果がユーザの端末に一覧表示される(図3、ステップSa2参照)。また、この時、ステップSb1の判断結果がYESとなって、所定時間(例えば1~2分間)以内にユーザが検索結果として一覧表示されたページのタイトルまたはURLのうち、いずれかをクリックしたか否かを判断する(ステップSb2)。

【0045】ここで、所定時間以内に一覧表示されたページのタイトルまたはURLが何等クリックされなかった場合は、判断結果がNOとなってマイクロユーザ属性推定部6は処理を終了する。また、所定時間以内にいずれ

かのタイトルまたはURLがクリックされた場合は、判断結果がYESとなり、マイクロユーザ行動ログ記憶部9に記憶されている上記ユーザのユーザ識別子に対応づけて、上記ユーザがステップSb1で入力された検索キーと、上記ユーザがクリックしたページのタイトルまたはURLとを、マイクロユーザ行動ログとしてマイクロユーザ行動ログ記憶部9に記憶する(ステップSb3)。ここで、上記累計は傾向の大きさ(トレンド強度)を示していると思なせる。

【0046】次に、そのクリックされたタイトルまたはURL(検索結果)が、マクロトレンド抽出部5によって抽出されているマクロトレンドに一致するか否かを判断する(ステップSb4)。ここで、この判断は、クリックされた検索結果のURLで示されるホームページに、マクロトレンドを示す単語が含まれているか、または、クリックされた検索結果のURL自体がマクロトレンドを決定する際に、他のユーザがクリックしている場合に、一致していると判断する。

【0047】そして、クリックされた検索結果が抽出されているマクロトレンドと一致しない場合は、判断結果がNOとなってマイクロユーザ属性推定部6は処理を終了する。一方、一致した場合は判断結果がYESとなって一致したマクロトレンドがユーザ属性を特定するトレンドか否かを判断する(ステップSb5)。ここで、ユーザ属性を特定するトレンドとは、ユーザの年齢、家族構成、居住地域、趣味、嗜好、購入したいもの、興味の対象、興味のあるイベント等、ユーザの属性を特定することが可能なトレンドのことをいう。例えば、ユーザの居住地域を特定するような場合、ユーザの居住地域に関する情報(例えば、ユーザの居住地域付近における桜の開花時期、開花予測等)をユーザが取得しようとした場合に、そのユーザの居住地域(属性)が特定される。その場合、ユーザ属性(居住地域)を特定したトレンドは、“居住地域付近における桜の開花時期または開花予測”ということになる。

【0048】もし、ユーザ属性を特定するトレンドであれば、判断結果がYESとなって、そのマクロトレンドをユーザ属性として記憶する(ステップSb6)。一方、ユーザ属性を特定するマクロトレンドでなければ、判断結果がNOとなってそのマクロトレンドを、ユーザ識別子と共に、ユーザ属性として記憶する(ステップSb7)。このステップSb6の処理により、今までマクロトレンドとしては確定されていなかったものが、新たなマクロトレンドとして確定される。

【0049】次に図2に戻り、ステップS7へ進み、マクロトレンドリンク付け部7では、同一の嗜好をもつマクロトレンド同士をリンク付けする。これは、物理的に同一の商品や、概念を別の言葉で表現しているものに対して、マクロトレンドに基づいてリンク付けすることにより、関係付けすることができる。

【0050】本実施形態のマクロトレンドリンク付け部7では、マクロトレンド抽出部5が推定したユーザのマクロトレンドに対して形態素解析を行う。形態素解析とは、入力された文字列を単語辞書に対して、検索を行い、品詞情報(品詞)、文頭可否情報(文頭可)、前方接続情報(前接)、後方接続情報(後接)等の情報を取得する。通常の単語辞書では、TREI辞書構造という特別な辞書構造を持つことにより、高速な検索を行えるようになっている。

【0051】例えば、“ああ”、“あいさつ”、“あい”という辞書項目がある場合、それぞれの第一文字(ここでは、日本語であるので、C言語の文字であるアルファベットと異なり、日本語文字2byteを示す)が同じもの、第二文字目が同じものなど、それぞれ順次木構造的に構成される。そして、最後の文字まで、一致した場合には、その単語辞書項目に対する品詞情報(品詞)、文頭可否情報(文頭可)、前方接続情報(前接)、後方接続情報(後接)等の情報が記述される。

【0052】ここで、文頭可否情報とは、文頭にあってよいかどうかを示すフラグである。文頭可であれば、文頭に存在してもよいが、文頭否であれば、文頭にあることが許可されない単語ということになる。また、前方接続情報とは、前の単語の品詞または属性が適正な場合だけ接続が許可され、前接で接続が許可されない単語の場合、候補として削除される。同様に、後方接続情報も、後の単語の品詞または属性が適正な場合だけ接続が許可され、後接で接続が許可されない単語の場合、候補として削除される。

【0053】このような品詞接続により最尤候補を選択する。最尤候補は、最小コスト法と呼ばれる方法により選択する。最小コスト法とは、最もコストが最小となる形態素候補を最尤候補とする処理方式である。形態素解析において利用されるコストには、一般に接続コストと単語コストの2種類のコストがある。

【0054】接続コストは、ある単語と単語を接続する場合に必要なコストである。単語と単語であるため、単語にその単語の活用を加えたものに対する接続コストは「0」となる。また、単語コストとは、その単語に関するコストであり、例えば、使用頻度が高い単語はコストが低くなる。また、活用は単語ではないので、コストは「0」となる。

【0055】この形態素解析により、テキスト部が単語単位に分解されると同時に、各単語に最も相応しいと考えられる品詞が付与される。例えば、“プレステ”、“ファージ”は、それぞれ図7(a)、(b)に示すように形態素解析される。

【0056】ここで、マクロトレンド抽出部5によって抽出されたマクロトレンドの中で、未知語または固有名詞を抽出する。それ以外のマクロトレンドは除く。そして、同時期または時間的に少しずれて発生したマクロ

レンドの中で、異なる表現のマクロトレンドを抽出する。そして、時間的に類似、または、マクロトレンドの傾向が類似する場合、そのマクロトレンドは、類似する対象（例えば商品名や事象等）を指し示すとして、マクロトレンドをリンク付けすることができる。ここで、時間的に類似とは、例えばある商品が日米同時発売されたり、また、日米で同日に発生したトレンドであれば、米国との時差分だけずれて同様なトレンドが発生するが、そのような場合をいう。

【0057】例えば、図8に示すように、米国でのマクロトレンドと日本でのマクロトレンドとを比較し、異なる表現で、かつ、傾向（時間の経過に伴うトレンド強度の変化）が類似するマクロトレンドとして、米国において「Furby」と表現されたマクロトレンドと、日本において「ファービ」と表現されたマクロトレンドとが抽出されたとすると、これらのマクロトレンドが対象とするものが類似の商品を指し示しているものとし、「Furby＝ファービ」として相互のマクロトレンドをリンク付けすることができる。

【0058】次に、ステップS8へ進み、ユーザ識別子収集部8では、各サイトからユーザ識別子のKeyを取り出し、そのKeyを基に各ユーザ識別子を集集する。各サイトに対してHTTPのset_cookie情報を受け、各サイトにあるKey情報を抽出し、そのKey情報をもとに、ある特定のURLサイトに入ってきたユーザに対して当該Keyをもとに、get_cookieすることにより、各ユーザ毎に、図1のユーザ識別子発生部2が発生したKeyおよびそのユーザ識別子と、他のサイトで提供しているKeyおよびそのユーザ識別子とを対応付け、ユーザ識別子テーブル記憶部11へ記憶させておく。そして、ユーザ識別子テーブル記憶部11からマイクロユーザ行動ログ記憶部9へ、ある個人ユーザが持つ複数のユーザ識別子を全て出力する。

【0059】例えば、ある個人ユーザUが、種々のサイトに訪問している場合、ユーザUは、特定のユーザ識別子のみを持つわけではなく、あるAサイトでは“1111”というユーザ識別子を、また、他のBサイトでは“2222”というユーザ識別子を持つ場合がある。よって、それらのユーザ識別子を統合するために、ユーザUのユーザ識別子が“1111”であり、かつ、“2222”であるという情報をユーザ識別子テーブル記憶部11に記憶させておき、ユーザ識別子テーブル記憶部11からマイクロユーザ行動ログ記憶部9へ、ある個人ユーザが持つ複数のユーザ識別子を全て出力する。

【0060】例えば、マイクロユーザ行動ログ記憶部9内に、ユーザUのユーザ識別子“1111”がURL: www.123.co.jpへジャンプしているマイクロユーザ行動ログと、同じくユーザUのユーザ識別子“2222”がURL: www.aaa.co.jpへジャンプしているマイクロユーザ行動ログが記憶されていた場合、それら2つのログを統合

し、ユーザUが、URL: www.123.co.jp およびURL: www.aaa.co.jpの2つの行動を行ったことを明示する。これにより、別のサイトで別のcookieを持っているユーザでも本実施形態のユーザ属性特定装置が提供するユーザ識別子で一元的に管理することができる。

【0061】なお、図1に示すインターフェイス部1、ユーザ識別子発生部2、ユーザ識別子送信部3、ユーザ識別子受信部4、マクロトレンド抽出部5、マイクロユーザ属性推定部6、マクロトレンドリンク付け部7、ユーザ識別子収集部8、マイクロユーザ行動ログ記憶部9、マイクロユーザ行動ログ記憶部10、ユーザ識別子テーブル記憶部11の機能を実現するためのプログラム、または、図2、図3、図6の手順を実現するためのプログラムを、コンピュータ読み取り可能な記録媒体に記録し、この記録媒体に記録されたプログラムをコンピュータシステムに読み込ませ実行することにより、ユーザに割り当てるネットワークの通信レートを制御するようにしてもよい。

【0062】ここで、上記「コンピュータシステム」とは、OSや周辺機器等のハードウェアを含み、さらにWWWシステムを利用している場合であれば、ホームページ提供環境（あるいは表示環境）も含むものとする。また、「コンピュータ読み取り可能な記録媒体」とは、フロッピーディスク、光磁気ディスク、ROM、CD-ROM等の可搬媒体、コンピュータシステムに内蔵されるハードディスク等の記憶装置のことをいう。さらに「コンピュータ読み取り可能な記録媒体」とは、インターネット等のネットワークや電話回線等の通信回線を介してプログラムを送信する場合の通信線のように、短時間の間、動的にプログラムを保持するもの（伝送媒体もしくは伝送波）、その場合のサーバやクライアントとなるコンピュータシステム内部の揮発性メモリのように、一定時間プログラムを保持しているものも含むものとする。

【0063】また、上記プログラムは、前述した機能の一部を実現するためのものであっても良く、さらに前述した機能をコンピュータシステムに既に記録されているプログラムとの組み合わせで実現できるものであっても良い。

【0064】

【発明の効果】以上説明したように、本発明によれば、目的付けられた行動に基づいて、ユーザの属性を自動的に推定することにより、ユーザの物理的に属する地域、地位、民族性、職業、性別、趣味などのユーザを特徴づける属性を自動的に抽出することができ、ユーザが意図的に記述したユーザ特性と異なり、ユーザの無意識下におけるユーザ特性を自動的に収集でき、マーケティングなどの基礎データとして利用することが可能となる。

【図面の簡単な説明】

【図1】 本発明に係るユーザ属性特定装置の一実施形態の構成を示すブロック図である。

【図2】 同ユーザ属性特定装置の動作を示すフローチャートである。

【図3】 同ユーザ属性特定装置におけるマクロトレンドの抽出手順を示すフローチャートである。

【図4】 マクロトレンドの一例を示す図である。

【図5】 図4に示したマクロトレンドを、地域毎に分けて抽出した場合を示す図である。

【図6】 本発明に係るユーザ属性特定装置におけるユーザ属性の特定手順を示すフローチャートである。

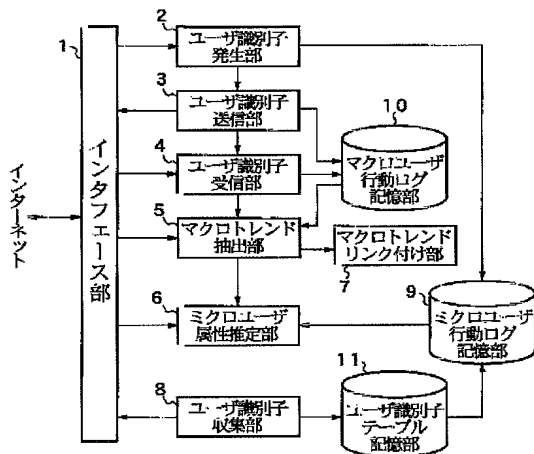
【図7】 形態素解析の一例を示す図である。

【図8】 マクロトレンドのリンク付けを説明するための説明図である。

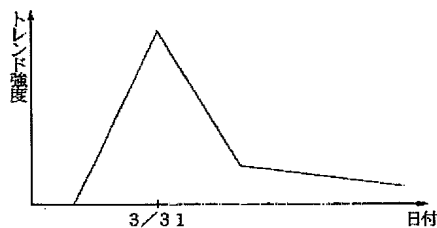
【符号の説明】

- 1 インターフェイス部
- 2 ユーザ識別子発生部
- 3 ユーザ識別子送信部
- 4 ユーザ識別子受信部
- 5 マクロトレンド抽出部
- 6 ミクロユーザ属性推定部
- 7 マクロトレンドリンク付け部
- 8 ユーザ識別子収集部
- 9 ミクロユーザ行動ログ記憶部
- 10 マクロユーザ行動ログ記憶部
- 11 ユーザ識別子テーブル記憶部

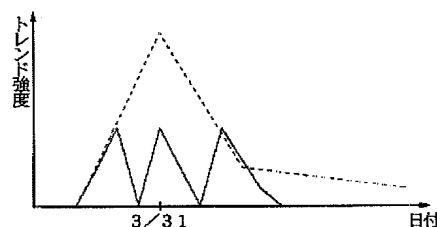
【図1】



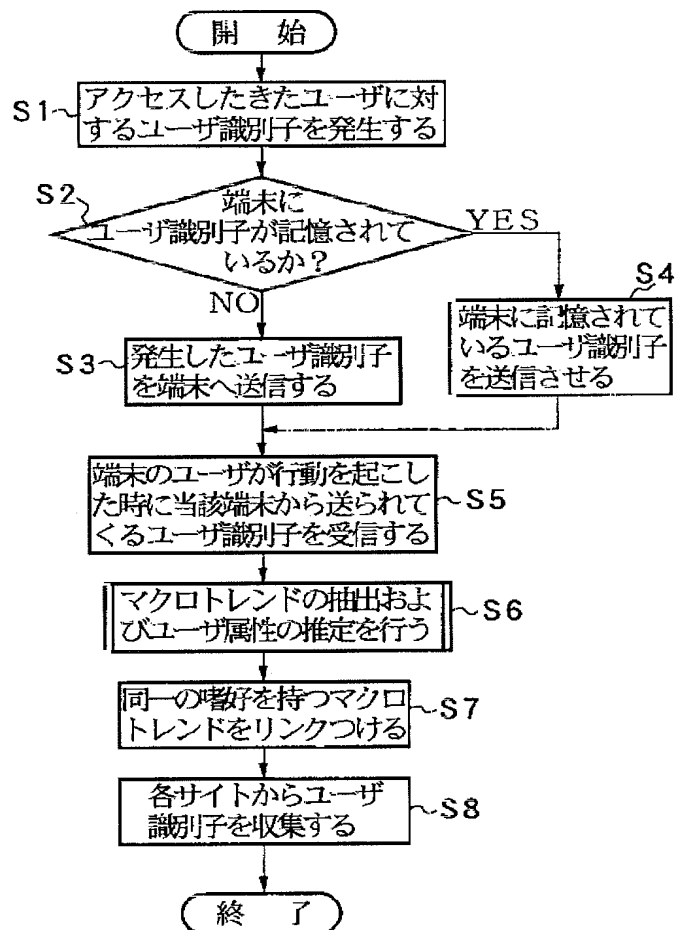
【図4】



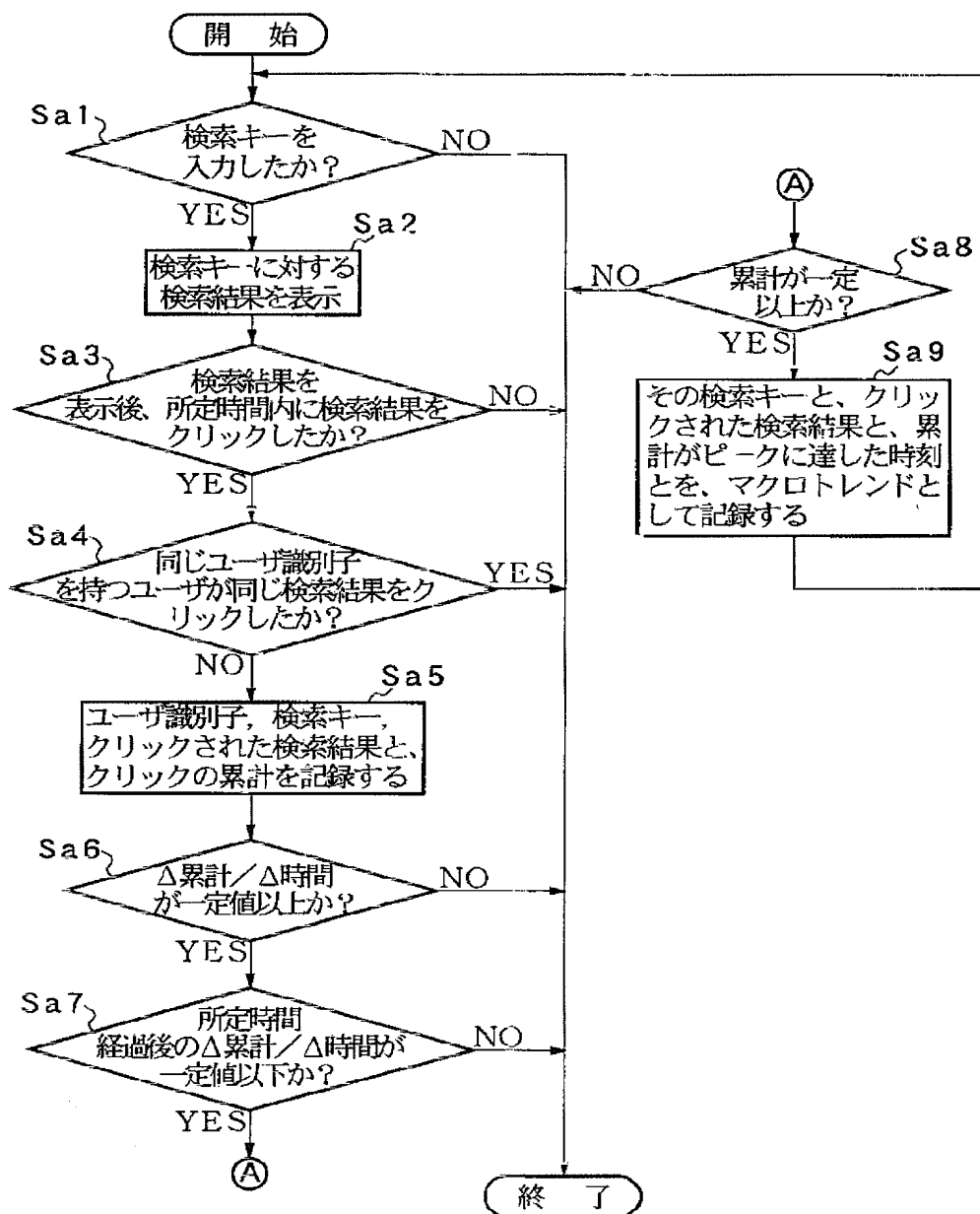
【図5】



【図2】



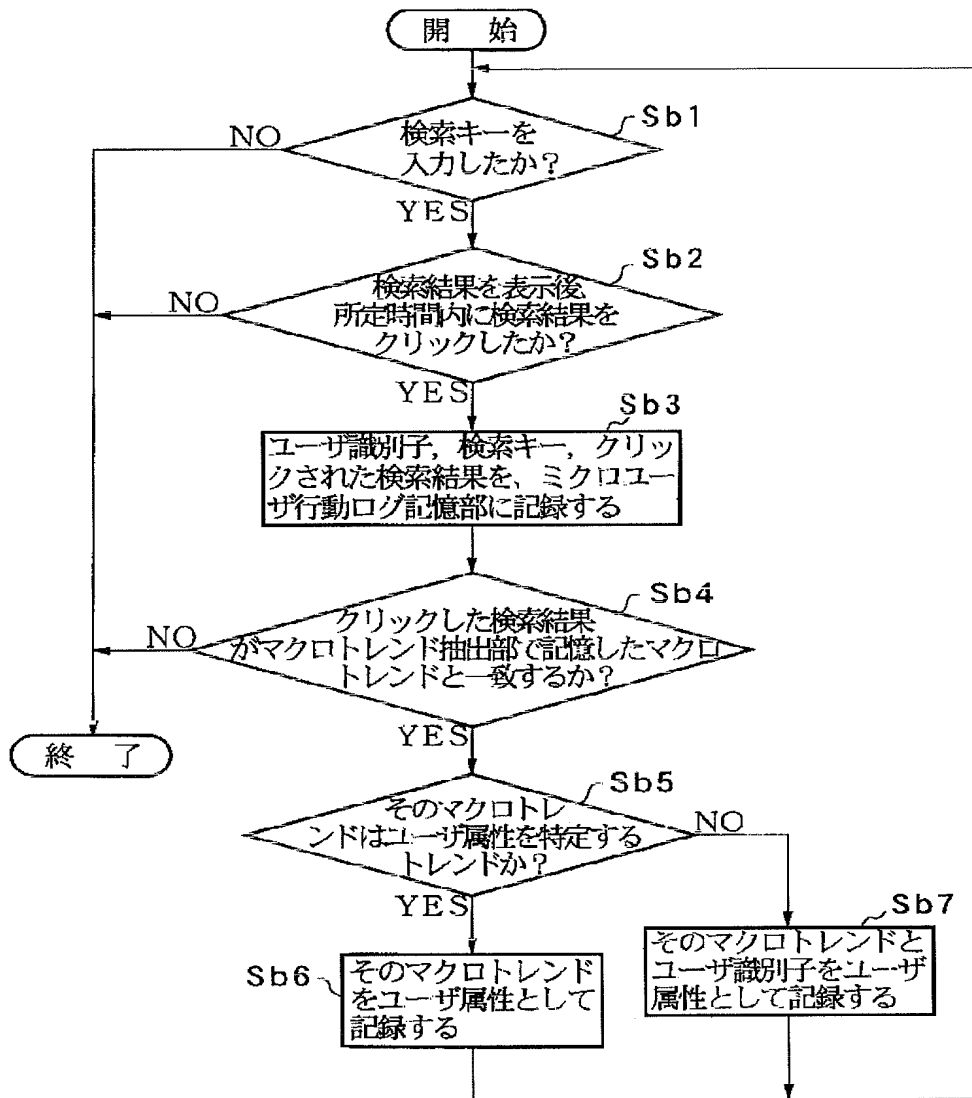
【図3】



【図8】



【図6】



【図7】

(a)

表記	品詞	未知語
プレイステーション		
cost=33		
プレイステーション	固有名詞	

(b)

表記	品詞	未知語
ファーマ		
cost=33		
ファーマ	未知語	○

フロントページの続き

Fターム(参考) 5B049 AA01 EE05 FF01 GG09
5B075 PR03 QM05
5B089 GA11 GB03 HA10 JA21 JA37
KA14 KC48 KC53